The Mean-Median-Mode Inequality and
χ^2 Distributions

by

Pranab K. Sen

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1818

February 1987

The Mean-Median-Mode Inequality and Noncentral χ^2 Distributions

BY PRANAB KUMAR SEN

University of North Carolina, Chapel Hill

SUMMARY

The usual ordering of the mean, median and mode for a central chi square distribution is extended for the noncentral case. In this context, a subadditive property of the median and mode is also established.

1. INTRODUCTION

For a distribution F with a density function f, defined on the real line R, the mean (μ), median (m) and mode (M) are defined by

$$\mu = \int_R x\,dF(x), \quad m = \inf\{x : F(x) \geq \tfrac{1}{2}\} \text{ and } M : f(M) = \sup_{x \in R} f(x), \tag{1.1}$$

where, we assume that M is unique. Pearson (1895) found empirically that for his Type III distributions,

$$\mu - M \quad 3(\mu - m), \tag{1.2}$$

which led him to the conjecture that for a distribution F,

$$\text{either } \mu \geq m \geq M \quad \text{or} \quad \mu \leq m \leq M, \tag{1.3}$$

and this is classically known as the mean-median-mode inequality. Although (1.2) may not generally hold for all distributions, the ordering in (1.3) remains of considerable interest, particularly, in the context of measuring skewness of a distribution (see MacGillivray (1985) and the references cited there). Though some sufficient conditions ensuring (1.3) have been laid down by van Zwet (1979) and MacGillivray (1981), among others, their verification may generally demand considerable manipulations.

A class of distributions arising frequently in problems of statistical inference is the family of (central and noncentral) chi square distributions.

For a central chi square distribution with p degrees of freedom (DF), denoting these measures by $\mu_p^{(0)}$, $m_p^{(0)}$ and $M_p^{(0)}$, we have

$$\mu_p^{(0)} = p,\ M_p^{(0)} = (p-2)\vee 0 \text{ and } p-1 < m_p^{(0)} < p, \tag{1.4}$$

so that (1.3) holds [see Johnson and Kotz (1970, ch. 17)]; actually, for $p \geq 2$, $m_p^{(0)}$ varies between $p - 1 + \cdot 386$ to $p - 1 + \cdot 34$, as p increases, so that (1.2) holds. The situation is more complicated for the noncentral case. For a noncentral chi square distribution with p DF and noncentrality parameter $\lambda (\geq 0)$, the mean is given by

$$\mu_p^{(\lambda)} = p + \lambda = \mu_p^{(0)} + \lambda, \tag{1.5}$$

so that $\mu_p^{(\lambda)}$ is additive with the noncentrality parameter. Though various approximations for the median $(m_p^{(\lambda)})$ have been discussed in Johnson and Kotz (1970, ch. 28), it is not known whether $m_p^{(\lambda)}$ is additive or not. Further, it is not precisely known whether the mode $(M_p^{(\lambda)})$, $m_p^{(\lambda)}$ and $\mu_p^{(\lambda)}$ have the same ordering as in the central case. The object of the present note is to focus on (1.3) for the noncentral chi square distributions, and to show that

$$m_p^{(\lambda)} \leq m_p^{(0)} + \lambda \text{ and } M_p^{(\lambda)} \leq M_p^{(0)} + \lambda, \forall \lambda \geq 0, \tag{1.6}$$

so that the mode and median are both sub-additive with respect to the noncentrality parameter. The main results are presented in the next section.

2. THE MAIN RESULTS

We denote the central chi square distribution function with p DF by $G_p(x)$, $x \geq 0$, $p \geq 1$, and the noncentral chi square distribution function with p DF and noncentrality parameter λ by $G_p^{(\lambda)}(x)$. Also, let

$$q_\lambda(r) = e^{-\frac{1}{2}\lambda}(\tfrac{1}{2}\lambda)^r/r!, \text{for } r \geq 0. \tag{2.1}$$

Then, it is well known that

$$G_p^{(\lambda)}(x) = \Sigma_{r\geq 0}\, q_\lambda(r) G_{p+2r}(x), \quad x \geq 0,\ \lambda \geq 0. \tag{2.2}$$

We also denote the density functions corresponding to $G_p(x)$ and $G_p^{(\lambda)}(x)$ by $g_p(x)$ and $g_p^{(\lambda)}(x)$, respectively, so that

$$g_p^{(\lambda)}(x) = \Sigma_{r\geq 0}\, q_\lambda(r) g_{p+2r}(x), \quad x \geq 0. \tag{2.3}$$

We write $\bar{G}_p(x) = 1 - G_p(x)$ and $\bar{G}_p^{(\lambda)}(x) = 1 - G_p^{(\lambda)}(x)$. Then, note that

$$G_{p-2}(x) - G_p(x) = 2g_p(x), \ \forall\, p \geq 2,\ x \geq 0, \tag{2.4}$$

so that by (2.2) and (2.4), we have

$$G_{p-2}^{(\lambda)}(x) = G_p^{(\lambda)}(x) + 2g_p^{(\lambda)}(x) \geq G_p^{(\lambda)}(x), \quad x \geq 0. \tag{2.5}$$

Using (2.5) and (2.4), it is easy to show that for every $p \geq 2$,

$$G_p^{(\lambda)} \text{ is unimodal (with mode } M_p^{(\lambda)} \geq 0). \tag{2.6}$$

Then, we have the following.

Theorem 1. For every $p(\geq 2)$ and $\lambda(\geq 0)$,

$$g_p^{(\lambda)}(M_p^{(\lambda)}) = g_{p-2}^{(\lambda)}(M_p^{(\lambda)}), \tag{2.7}$$

$$M_p^{(\lambda)} \leq p - 2 + \lambda = M_p^{(0)} + \lambda, \ \forall\, \lambda \geq 0, \tag{2.8}$$

$$M_p^{(\lambda)} \geq p - 2 + p^{-1}(p-2)\lambda, \ \forall\, \lambda \geq 0. \tag{2.9}$$

Proof. (2.7) is well-known [viz., Johnson and Kotz (1970, ch. 28, p. 136)]. Nevertheless, it provides the key to the other assertions. First, by (2.3),

$$\begin{aligned}(\partial/\partial x)g_p^{(\lambda)}(x) &= \Sigma_{r\geq 0} q_\lambda(r)(d/dx)g_{p+2r}(x) \\ &= \Sigma_{r\geq 0} q_\lambda(r)\{-\tfrac{1}{2}g_{p+2r}(x) + \tfrac{1}{2}g_{p+2r-2}(x)\} \\ &= -\tfrac{1}{2}\Sigma_{r\geq 0} q_\lambda(r)g_{p+2r}(x) + \tfrac{1}{2}\Sigma_{r\geq 0} q_\lambda(r)g_{p-2+2r}(x) \\ &= [g_{p-2}^{(\lambda)}(x) - g_p^{(\lambda)}(x)]/2, \quad \forall\, x \geq 0. \end{aligned} \tag{2.10}$$

Hence, by the unimodality in (2.6) and (2.10), (2.7) follows directly. Next, we note that (2.10) is positive or negative according as x is $<$ or $>$ $M_p^{(\lambda)}$, while $g_p^{(\lambda)}(M_p^{(\lambda)}) \geq g_p^{(\lambda)}(x)$, $\forall\, x \neq M_p^{(\lambda)}$. This implies that for $p \geq 2$,

$$g_{p-2}^{(\lambda)}(M_{p-2}^{(\lambda)}) \geq g_p^{(\lambda)}(M_p^{(\lambda)}) \text{ and } M_{p-2}^{(\lambda)} \leq M_p^{(\lambda)}, \ \forall\, \lambda \geq 0. \tag{2.11}$$

Next, we note that by (2.3), (2.7) and (2.10),

$$\Sigma_{r\geq 0} q_\lambda(r)g_{p+2r}(M_p^{(\lambda)})\{1 - (p + 2r - 2)/M_p^{(\lambda)}\} = 0, \tag{2.12}$$

and this, in turn, leads us to the following implicit equation for $M_p^{(\lambda)}$:

$$M_p^{(\lambda)} = \{\Sigma_{r\geq 0} q_\lambda(r)(p - 2 + 2r)g_{p+2r}(M_p^{(\lambda)})\}/g_p^{(\lambda)}(M_p^{(\lambda)})$$

$$= (p-2) + \{\Sigma_{r\geq 0} 2rq_\lambda(r) g_{p+2r}(M_p^{(\lambda)})\}/g_p^{(\lambda)}(M_p^{(\lambda)})$$

$$= (p-2) + \lambda\{\Sigma_{r\geq 0} q_\lambda(r) g_{p+2+2r}(M_p^{(\lambda)})\}/g_p^{(\lambda)}(M_p^{(\lambda)})$$

$$= (p-2) + \lambda g_{p+2}^{(\lambda)}(M_p^{(\lambda)})/g_p^{(\lambda)}(M_p^{(\lambda)}). \tag{2.13}$$

Making use of (2.11) (with p replaced by p+2), we obtain that

$$g_{p+2}^{(\lambda)}(m)/g_p^{(\lambda)}(m) \leq 1, \ \forall m \leq M_{p+2}^{(\lambda)} (\geq M_p^{(\lambda)}), \tag{2.14}$$

and hence, (2.8) follows directly from (2.13) and (2.14).

Next, we note that for every $m \geq 0$,

$$g_{p+2}^{(\lambda)}(m)/g_p^{(\lambda)}(m) = m\{\Sigma_{r\geq 0} q_\lambda(r) g_{p+2r}(m)(p+2r)^{-1}\}/g_p^{(\lambda)}(m)$$

$$\geq m\{\Sigma_{r\geq 0} q_\lambda(r) g_{p+2r}(m)(p+2r)/g_p^{(\lambda)}(m)\}^{-1}$$

$$= m\{p + \lambda g_{p+2}^{(\lambda)}(m)/g_p^{(\lambda)}(m)\}^{-1}. \tag{2.15}$$

From (2.15), we readily obtain that

$$g_{p+2}^{(\lambda)}(M_p^{(\lambda)})/g_p^{(\lambda)}(M_p^{(\lambda)}) \geq M_p^{(\lambda)}/(p + \lambda), \tag{2.16}$$

so that by (2.13) and (2.16), we have

$$M_p^{(\lambda)} \geq (p-2) + \lambda M_p^{(\lambda)}/(p + \lambda), \tag{2.17}$$

and (2.9) follows readily from (2.17). This completes the proof of the Theorem.

It may be noted that (2.9) is not sharp for $\lambda > 0$ (and particularly, for large λ). It is possible to exploit (2.15) to derive better (lower) bounds for $M_p^{(\lambda)}$ for λ large. In fact, we may easily verify that by (2.13) and (2.15), for every $\lambda \geq 0$, $M_p^{(\lambda)}$ satisfies the following inequality:

$$M_p^{(\lambda)}\{1 - 2\lambda/[p + \sqrt{p^2 + 4\lambda M_p^{(\lambda)}}\,]\} \geq (p-2). \tag{2.18}$$

Next, we consider another inequality on the mode which will be very useful for the study of the median inequalities. Note that by (2.7),

$$g_{p+2}^{(\lambda)}(M_{p+2}^{(\lambda)}) = g_p^{(\lambda)}(M_{p+2}^{(\lambda)}), \tag{2.19}$$

so that proceeding as in (2.13), we obtain that

$$M_{p+2}^{(\lambda)} - (p-2) = \{\Sigma_{r\geq 0} q_\lambda(r) g_{p+2r+2}(M_{p+2}^{(\lambda)})(2r+2)\}/g_{p+2}^{(\lambda)}(M_{p+2}^{(\lambda)})$$

$$\geq \{\Sigma_{r\geq 0}(2r+2)^{-1} q_\lambda(r) g_{p+2r+2}(M_{p+2}^{(\lambda)})/g_{p+2}^{(\lambda)}(M_{p+2}^{(\lambda)})\}^{-1}$$

$$= \{\tfrac{1}{\lambda}\Sigma_{r\geq 0} q_\lambda(r+1) g_{p+2(r+1)}(M_{p+2}^{(\lambda)})/g_{p+2}^{(\lambda)}(M_{p+2}^{(\lambda)})\}^{-1}$$

$$= \lambda\{-q_\lambda(0) g_p(M_{p+2}^{(\lambda)})/g_{p+2}^{(\lambda)}(M_{p+2}^{(\lambda)}) + g_p^{(\lambda)}(M_{p+2}^{(\lambda)})/g_{p+2}^{(\lambda)}(M_{p+2}^{(\lambda)})\}^{-1}$$

$$= \lambda g_{p+2}^{(\lambda)}(M_{p+2}^{(\lambda)})/\{g_p^{(\lambda)}(M_{p+2}^{(\lambda)}) - q_\lambda(0) g_p(M_{p+2}^{(\lambda)})\}$$

$$= \lambda g_p^{(\lambda)}(M_{p+2}^{(\lambda)})/\{g_p^{(\lambda)}(M_{p+2}^{(\lambda)}) - q_\lambda(0) g_p(M_{p+2}^{(\lambda)})\} \text{ [by (2.19)]}$$

$$\geq \lambda, \ \forall \ p, \lambda. \tag{2.20}$$

Therefore, we obtain that

$$M_{p+2}^{(\lambda)} > p - 2 + \lambda, \ \forall \ p, \lambda. \tag{2.21}$$

Next, we consider the following theorem on the median.

<u>Theorem 2</u>. <u>For any fixed</u> p, $m_p^{(\lambda)}$ <u>is</u>   <u>in</u> $\lambda(\geq 0)$, $m_p^{(\lambda)} \geq m_{p-2}^{(\lambda)}$, <u>for every</u> $p \geq 2$ <u>and</u> $\lambda \geq 0$, <u>and for every</u> $p \geq 2$, $\lambda \geq 0$

$$M_p^{(\lambda)} \leq p - 2 + \lambda < m_p^{(\lambda)} \leq p + \lambda = \mu_p^{(\lambda)}. \tag{2.22}$$

<u>Proof</u>. Note that $\bar{G}_p(x)$ is ↘ in x. Also, note that

$$(\partial/\partial\lambda)\bar{G}_p^{(\lambda)}(p + \lambda) = (\partial/\partial\lambda)\{\Sigma_{r\geq 0} q_\lambda(r)\bar{G}_{p+2r}(p + \lambda)\}$$

$$= \{\Sigma_{r\geq 0}(-\tfrac{1}{2}q_\lambda(r) + \tfrac{1}{2}q_\lambda(r-1)\}\bar{G}_{p+2r}(p + \lambda)) - \Sigma_{r\geq 0} q_\lambda(r) g_{p+2r}(p + \lambda)\}$$

$$= [\bar{G}_{p+2}^{(\lambda)}(p + \lambda) - \bar{G}_p^{(\lambda)}(p + \lambda)]/2 - g_p^{(\lambda)}(p + \lambda)$$

$$= g_{p+2}^{(\lambda)}(p + \lambda) - g_p^{(\lambda)}(p + \lambda). \tag{2.23}$$

Note that by (2.8), $M_{p+2}^{(\lambda)} \leq p + \lambda$, $\forall \ \lambda \geq 0$, while in (2.14), the opposite inequality holds for $m \geq M_{p+2}^{(\lambda)}$. Hence, the right hand side of (2.23) is always non-negative. Therefore, for every $\lambda \geq 0$,

$$\bar{G}_p^{(\lambda)}(p + \lambda) - \bar{G}_p^{(0)}(p + 0) = \int_0^\lambda \{(\partial/\partial\alpha)\bar{G}_p^{(\alpha)}(p + \alpha)\}dx \geq 0. \tag{2.24}$$

On the other hand, $\bar{G}_p^{(0)}(p + 0) = \bar{G}_p(p) < \bar{G}\ (m_p^{(0)}) = \tfrac{1}{2}$, $\forall \ p \geq 1$. Therefore, we conclude that

$$\bar{G}_p^{(\lambda)}(p + \lambda) \text{ is monotone increasing in } \lambda(\geq 0). \tag{2.25}$$

Further, note that

$$\lim_{\lambda \to \infty} G_p^{(\lambda)}(p + \lambda) = \lim_{\lambda \to \infty} P\{\chi^2_{p,\lambda} \geq p + \lambda\}$$

$$= \lim_{\lambda \to \infty} P\{(\chi^2_{p,\lambda} - (p + \lambda))/(2(p + 2\lambda))^{\frac{1}{2}} \geq 0\}$$

$$= 1/2, \tag{2.26}$$

where the last step follows from the asymptotic normality results on the non-central chi square statistic; it is also possible to use other refined approximations, discussed in Chapter 28 of Johnson and Kotz (1970). By (2.25) and (2.26), we conclude that for every p,

$$\bar{G}_p^{(\lambda)}(p + \lambda) \leq \tfrac{1}{2} \text{ for every finite } \lambda (\geq 0). \tag{2.27}$$

On the other hand, by definition,

$$\bar{G}_p^{(\lambda)}(m_p^{(\lambda)}) = \tfrac{1}{2}, \ \forall \lambda \geq 0, \ p \geq 1, \tag{2.28}$$

so that noting that $\bar{G}_p^{(\lambda)}(x)$ is $\searrow$ in x, we immediately conclude from (2.27) and (2.28) that

$$m_p^{(\lambda)} \leq p + \lambda, \ \forall \lambda \geq 0. \tag{2.29}$$

Next, we consider (2.23) at the point $p - 2 + \lambda$, and get that

$$(\partial/\partial\lambda)\bar{G}_p^{(\lambda)}(p - 2 + \lambda) = g_{p+2}^{(\lambda)}(p - 2 + \lambda) - g_p^{(\lambda)}(p - 2 + \lambda). \tag{2.30}$$

Note that by (2.21), $p - 2 + \lambda \leq M_{p+2}^{(\lambda)}$, so that by (2.14) and (2.30), we obtain that

$$(\partial/\partial\lambda)\bar{G}_p^{(\lambda)}(p - 2 + \lambda) \text{ is } \leq 0, \ \forall \lambda \geq 0, \tag{2.31}$$

so that parallel to (2.25), we obtain that

$$\bar{G}_p^{(\lambda)}(p - 2 + \lambda) \text{ is } \searrow \text{ in } \lambda (\geq 0); \ \bar{G}_p^{(0)}(p - 2) > \tfrac{1}{2}. \tag{2.32}$$

Again, proceeding as in (2.26), we have

$$\lim_{\lambda \to \infty} \bar{G}_p^{(\lambda)}(p - 2 + \lambda) = \tfrac{1}{2}, \tag{2.33}$$

so that

$$\bar{G}_p^{(\lambda)}(p - 2 + \lambda) \geq \tfrac{1}{2}, \ \forall \text{(finite)}\lambda(\geq 0). \tag{2.34}$$

This, along with (2.28), imply that

$$m_p^{(\lambda)} \geq p - 2 + \lambda, \ \forall \lambda \geq 0. \tag{2.35}$$

Since $M_p^{(\lambda)} \leq p - 2 + \lambda, \ \forall \lambda \geq 0$ [(2.8)], (2.22) follows directly from (2.29) and (2.35). Also, note that (2.21) and (2.22) imply that

$$m_{p+2}^{(\lambda)} \geq p + \lambda \geq m_p^{(\lambda)}, \ \forall p, \lambda \geq 0. \tag{2.36}$$

Further, $\bar{G}_q^{(\lambda)}(x)$ is $\nearrow$ in λ (for any given x and q), so that for $\lambda_2 \geq \lambda_1 \geq 0$,

$$\tfrac{1}{2} = \bar{G}_p^{(\lambda_1)}(m_p^{(\lambda_1)}) \leq \bar{G}_p^{(\lambda_2)}(m_p^{(\lambda_1)}); \ \bar{G}_p^{(\lambda_2)}(m_p^{(\lambda_2)}) = \tfrac{1}{2}, \tag{2.37}$$

hence $m_p^{(\lambda_2)} \geq m_p^{(\lambda_1)}, \ \forall \lambda_1 \leq \lambda_2$. This completes the proof of the theorem.

Note that by definition, $(\partial/\partial\lambda)\bar{G}_p^{(\lambda)}(m_p^{(\lambda)}) = 0$, and this yields that

$$\begin{aligned}(\partial/\partial\lambda)m_p^{(\lambda)} &= g_{p+2}^{(\lambda)}(m_p^{(\lambda)})/g_p^{(\lambda)}(m_p^{(\lambda)}) \\ &= 1 + \{g_{p+2}^{(\lambda)}(m_p^{(\lambda)}) - g_p^{(\lambda)}(m_p^{(\lambda)})\}/g_p^{(\lambda)}(m_p^{(\lambda)}) \\ &= 1 + 2\{(d/dm)\log g_p^{(\lambda)}(m)|_{m=m_p^{(\lambda)}}\},\end{aligned} \tag{2.38}$$

as $(d/dx)g_p^{(\lambda)}(x) = \{g_{p+2}^{(\lambda)}(x) - g_p^{(\lambda)}(x)\}/2, \ \forall x \geq 0$. Further, $g_p^{(\lambda)}(x)$ has a unique mode at $x = M_p^{(\lambda)}$ ($\leq m_p^{(\lambda)}$, by (2.22)), and hence, we have $(d|dm)\log g_p^{(\lambda)}(m)|_{m=m_p^{(\lambda)}} \leq 0, \ \forall \lambda \geq 0, \ p \geq 2$. Therefore, by (2.38), we claim that $(\partial/\partial\lambda)m_p^{(\lambda)} \leq 1, \ \forall \lambda \geq 0$, so that

$$m_p^{(\lambda)} - m_p^{(0)} \leq \lambda, \ \forall \lambda \geq 0, \tag{2.39}$$

and this verifies (1.6). Since $g_{p+2}^{(\lambda)}(m)/g_p^{(\lambda)}(m)$ is $\gtreqless 1$ according as m is $\gtreqless M_{p+2}^{(\lambda)}$, and as $(\partial/\partial\lambda)m_p^{(\lambda)} \leq 1$, we claim by using (2.38) that $m_p^{(\lambda)} \leq M_{p+2}^{(\lambda)}$, for every $\lambda \geq 0$. Thus, we may even strengthen (2.22) to the following:

$$M_p^{(\lambda)} \leq p - 2 + \lambda \leq m_p^{(\lambda)} \leq M_{p+2}^{(\lambda)} \leq p + \lambda = \mu_p^{(\lambda)}, \tag{2.40}$$

for all $\lambda \geq 0$. Also, while both $M_p^{(\lambda)}$ and $m_p^{(\lambda)}$ are subadditive (with respect to the noncentrality parameter λ), the mean $\mu_p^{(\lambda)}$ is additive.

We have considered so far the case of $p \geq 2$, althugh the results hold, in general, for all $p \geq 0$. For $p = 0$, $\chi^2_{p,\lambda} = \lambda$, with probability one, so that $\mu_0^{(\lambda)} = m_0^{(\lambda)} = M_0^{(\lambda)} = \lambda$, while for $p = 1$, the normal distribution can readily be called on to prove the mean-median-mode inequality. However, in this case, $M_1^{(\lambda)} = 0$ with $g_1^{(\lambda)}(0) = +\infty$, although $m_1^{(\lambda)} > 0$, $\lambda \geq 0$. For $p = 2$, $M_2^{(0)} = 0 < M_2^{(\lambda)}$, $\forall\ \lambda > 0$ and $g_2^{(\lambda)}(M_2^{(\lambda)}) < \infty$, $\forall\ \lambda > 0$. For Poisson mixtures of other gamma distributions too, parallel results hold. Finally, we may also remark that Haldane (1942) has remarked on the adequacy of (1.2) for the gamma family of densities. While his treatment may run into complications for the noncentral case, it appears that by virtue of (1.6) and (2.40) (and the fact that for the central case (1.2) holds approximately) that (1.2) holds even better for the noncentral case.

The author is grateful to Professor Norman L. Johnson for some valuable discussions on this topic. This work was supported by the National Institute of Health, Contract No. NO1-HV-12243-L.

REFERENCES

HALDANE, J.B.S. (1942). The mode and median of a nearly normal distribution with given cumulants. Biometrika 32, 294-299.

JOHNSON, N.L. AND KOTZ, S. (1970). Distributions in Statistics: Continuous Univariate Distributions, I, II. John Wiley, New York.

MACGILLIVRAY, H.L. (1981). The mean, median, mode inequality and skewness for a class of densities. Austral. J. Statist. 23, 247-250.

MACGILLVARY, H.L. (1985). The mean, median, mode and skewness. In Encyclo. Statist. Scs. 5, 364-367 (Eds. S. Kotz & N.L. Johnson, John Wiley, New York).

PEARSON, K. (1985). Contributions to the Mathematical Theory of evolution, II. Phil. Trans. Roy. Soc. London, 186, 343.

VAN ZWET, W.R. (1979). Mean, median, mode, II. Statist. Neerland 33, 1-5.